REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
09/13/02

2. REPORT DATE
Final Report

3. DATES COVERED *(From - To)*
06/01/95-09/30/01

4. TITLE AND SUBTITLE
A Dialogue-Based Architecture for a Tactical Picture Agent

5a. CONTRACT NUMBER
N00014-95-1-1008

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

6. AUTHOR(S)
Allen, James F.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Computer Science Dept.
University of Rochester Box 270226
Rochester, NY 14627-0226

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
800 N Quincy Ave.
Ballston Tower One
Arlington, VA 22217

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
Distribution of this document is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
This project developed technology to enable multi-modal dialogue-based interfaces for interactive analysis and problem solving. We developed a system that acts as a planning assistant, interacting with the user in conversational English to help with situation assessment and logistics planning. This enabled users without prior training to interact successfully with a set of sophisticated AI reasoning tools. The novel system architecture we developed will serve as the starting point for a new generation of mixed-initiative spoken dialogue-based systems.

15. SUBJECT TERMS
dialogue; intelligent interfaces

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPONE NUMBER *(Include area code)*
UU	UU	UU	UU	4	James F. Allen / 585-273-5218

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI-Std Z39-18

Final Report for
A Dialogue-based Architecture for a Tactical Picture Agent
ONR grant N00014-95-1-1008
ending Sept. 30, 2001

James F. Allen

Department of Computer Science
University of Rochester
Rochester, NY, USA
http://www.cs.rochester.edu/research/trains/

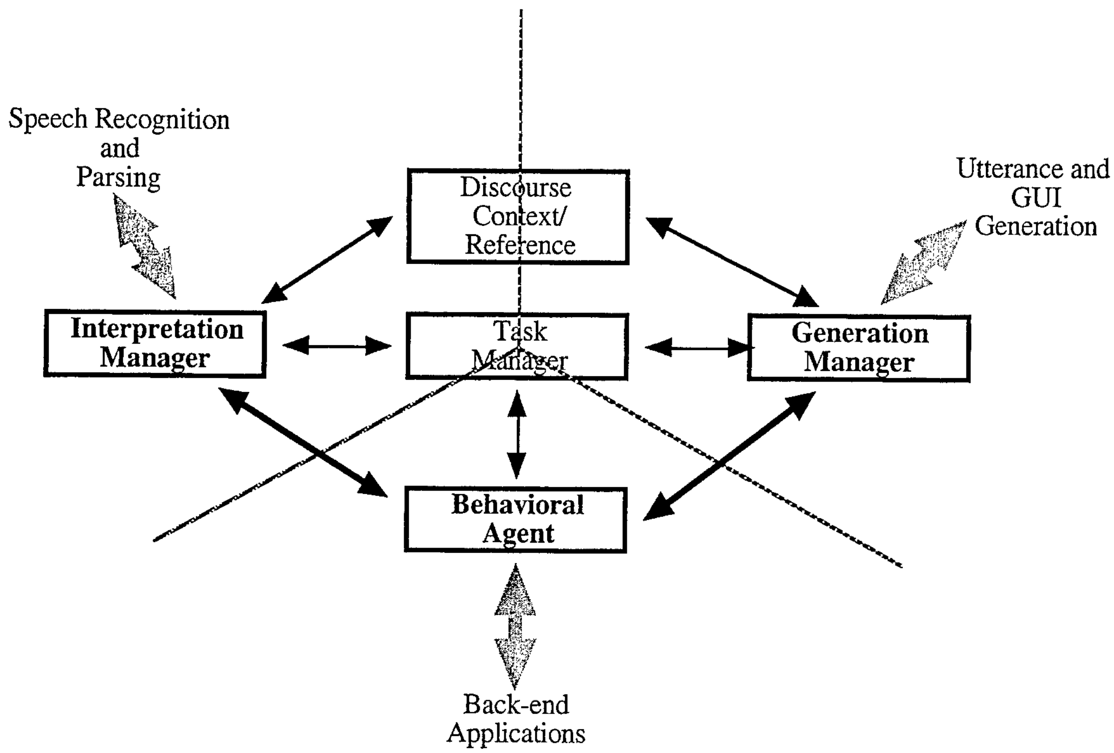
1. Executive Summary

This project developed technology to enable multi-modal dialogue-based interfaces for interactive analysis and problem solving. We developed a system that acts as a planning assistant, interacting with the user in conversational English to help with situation assessment and logistics planning. This approach enabled users without any prior training to interact successfully with a set of sophisticated AI reasoning tools including planners, schedulers, and knowledge based reasoning systems. The novel system architecture we developed will serve as the starting point for a new generation of mixed-initiative spoken dialogue-based systems.

2. Approach

One of the main premises in the project was that effective interactive situation assessment and collaborative plan development should have many of the characteristics of human dialogue, especially in the way topics are developed incrementally, then refined, clarified, and corrected iteratively until an acceptable solution is found. We developed a model where the human planner is always in the loop, and where both the human and the planning system could make contributions to the plans under development. The system recognizes the human planners intentions, determines how these affect the plan(s) under development, reasoned about the results of changes to the plans, and finally presented this information as a natural part of the dialogue. The state of the dialogue was captured by an abstract plan that serves as the interface between the humans perception of the goals and solutions so far, and the specialized domain reasoners (such as schedulers, routes planners, feasibility assessment, etc).

The research was driven by experimentation. We built robust end-to-end dialogue systems that interact with a user in order to accomplish a quantifiable task. By *robust*, we mean that everyday people could sit down at the machine and with less than a minute of instruction, engage in a dialogue to solve problems. By *end-to-end*, we mean that the system is complete: from speech input, through planning and reasoning and back to speech output. By a *quantifiable task* we mean that these were real problems for which there are right and wrong answers, or measurable degrees of effectiveness of solutions.



3. The Architecture

One of the key results of this project was the new system architecture. Figure 1 shows the key new parts of the architecture. The earlier version of the system had a centralized Discourse Manager that controlled the interpretation of user input and the planning of system responses. In the new design these functions are split apart. The Interpretation Manager (IM) is responsible for interpreting the user input and computing any discourse obligations (which constrain allowable system responses). The Behavioral Agent (BA) is responsible for determining the system's overall behavior, using problem solving obligations that it computes from users problem solving actions (computed by the IM) plus using information from other information sources as well. The Generation Manager (GM) is responsible for determining the system's communication with the user, and it uses the discourse obligations from the IM as well as the actions to perform from the BA. This new architecture has one key new technological advanatge and two key theoretical advantages:

1. It improves the portability of the system to new domains;
2. It allows the system to take more initiative in dialogues rather than mainly responding to the user input;
3. It allows the system to be more responsive and natural (e.g., the GM can provide acknowledgement even when the BA is deciding on a response).

In addition, we have made many advances in a range of topics relating to reference resolution and discourse context, generation, and robust parsing. These results are described in the papers listed below.

References

Allen, J.F., "Mixed-initiative interaction, trends and controversies," *IEEE Intelligent Systems*, September/October 1999.

Allen, J.F., "Natural language processing," *MIT Encyclopedia of Computer Science*. MIT Press, 1999.

Allen, J.F., D.K. Byron, M.O. Dzikovska, G.M. Ferguson, L. Galescu, and A.J. Stent, "An architecture for a generic dialogue shell," *J. Natural Language Engineering* (Special Issue on Best Practices in Spoken Language Dialogue Systems Eng.), 2000.

Allen, J.F., G.M. Ferguson, B.W. Miller, E.K. Ringger, and T.M. Sikorski (Zollo), "Dialogue systems: From theory to practice in TRAINS-96," in R. Dale, H. Moisl, and H. Somers (Eds.). *Handbook of Natural Language Processing*. New York: Marcel Dekkar, Inc., 347–376, 2000.

Blaylock, N.J., "Retroactive recognition of interleaved plans for natural language dialogue," TR 761, Computer Science Dept., U. Rochester, December 2001.

Blaylock, N.J., J.F. Allen, and G.M. Ferguson, "Managing communicative intentions with collaborative problem-solving," in R. Smith and J. van Kuppevelt (Eds). *Current and New Directions in Discourse and Dialogue*. Kluwer Academic Publishers, to appear, 2002.

Burstein, M., G.M. Ferguson, and J.F. Allen, "Integrating agent-based mixed-initiative control with an existing multi-agent planning system," *ICMAS-2000*, June 2000; extended version, TR 729, Computer Science Dept., U. Rochester, May 2000.

Byron, D.K., "Improving discourse management in TRIPS-98," *Proc., 6th European Conf. on Speech Communication and Technology* (Eurospeech-99), Budapest, September 1999.

Byron, D.K. and J.F. Allen, "Applying genetic algorithms to pronoun resolution," *Proc., 16th Nat'l. Conf. on Artificial Intelligence* (AAAI-99), Orlando, FL, July 1999.

Byron, D.K. and M. Dzikovska, "When is a union really an intersection? Problems resolving reference to locations in a dialogue system," *GÖTALOG Workshop on the Semantics and Pragmatics of Dialogue*, Göteborg, Sweden, June 2000.

Dzikovska, M.O., "A practical semantic type representation for natural language understanding," TR 733, Computer Science Dept., U. Rochester, August 2000.

Galescu, L. and E.K. Ringger, "Augmenting words with linguistic information for n-gram language models," *Proc., 6th European Conf. on Speech Communication and Technology* (Eurospeech-99), Budapest, September 1999.

Galescu, L. and J.F. Allen, "Bi-directional conversion between graphemes and phonemes using a joint n-gram model," *Proc., 4th ISCA Tutorial and Research Workshop on Speech Synthesis*, Perthshire, Scotland, August-September 2001.

Heeman, P.A. and J.F. Allen, "Improving robustness by modeling spontaneous speech events," in J.-C. Junqua and G. van Noord (Eds.). *Robustness in Language and Speech Technology*. Kluwer Academic Publishing, 167–196, 2000.

Heeman, P.A. and J.F. Allen, "Speech repairs, intonational phrases and discourse markers: Modeling speaker's utterances in spoken dialog," *Computational Linguistics 25*, 4, 527–572, 1999.

Lesh, N. and J.F. Allen, "Simulation-based inference for plan monitoring," *Proc., 16th Nat'l. Conf. on Artificial Intelligence* (AAAI-99), Orlando, FL, July 1999.

Ringger, E.K., "Correcting speech recognition errors," Ph.D. Thesis, Computer Science Dept., U. Rochester, April 2000.

Stent, A.J., "Content planning and generation in continuous-speech spoken dialog systems," *Proc., "May I Speak Freely?" Workshop* (to be held at the *German Annual Conf. on AI* (KI-99), Bonn, September 1999.

Stent, A.J., "The Monroe corpus," TR 728 and TN 99-2, Computer Science Dept., U. Rochester, March 2000.

Stent, A.J., "Rhetorical structure in dialog," *Proc., 2nd Int'l. Natural Language Generation Conf.* (INLG'2000), Mitzpe Ramon, Israel, June 2000.

Wolters, M. and D.K. Byron, "Functions of prosody for pronominal anaphora," *18th Int'l. Conf. on Computational Linguistics* (COLING-2000), Saarbrücken, July/August 2000.